

地理信息系统数据的不确定性问题

邬伦^①, 承继成^①, 史文中^②

(^① 北京大学遥感与地理信息系统研究所, 北京 100871; ^② 香港理工大学土地测量与地理资讯学系, 香港)

【摘要】 在总结当前 GIS 数据不确定性问题的研究进展和动态的基础上, 论述 GIS 数据不确定性的框架体系, 并分析探讨 GIS 数据不确定性的核心理论和主要研究内容, 特别是位置不确定性、属性不确定性、时域不确定性、不确定性传播和管理等问题, 最后对 GIS 数据不确定性的数学研究方法进行分析、归纳和阐述。

【关键词】 不确定性; 数据不确定性; 地理信息系统

【中图分类号】 P208

【文献标识码】 A

【文章编号】 1009-2307(2006)05-0013-05

1 引言

随着 GIS 研究和应用的深入, GIS 数据质量以及不确定性问题引起了广泛的关注。GIS 不确定性的研究内容与体系、数据不确定性的来源、数据不确定性处理方法、空间分析过程中的不确定性传播机理等已经成为 GIS 领域的研究热点。

不确定性是指客观世界或实体本身就具有的变异, 表现为不精确性、随机性和模糊性。由于人对客观实体和现象认识的局限性和表达的模糊性, 使原始数据本身存在不确定性, 以此数据进行 GIS 的分析与处理, 必然使分析结果也具有不确定性。

GIS 数据的不确定性一方面是由于客观世界存在的固有的复杂性、不稳定性, 和人们对其认识的不完备性; 另一方面, 在数据采集、录入、编辑、处理、分析和表述过程中出现的各种误差, 包括如系统误差或随机误差带来的不确定性。数据中的不确定性具有随机性、模糊性、灰色性、不完整性和未确知性等特征。GIS 数据的来源多种多样, 例如有的从调查与统计数据中得来, 有的则从地图和遥感数据中得来。其中遥感数据本身就含有不确定性的特征^[1], 比如“混合单元”问题。遥感数据虽然经过了各种校正处理(如光谱校正和几何校正), 但仍然保存了“残余(存)误差”, 这使得经过解译后产生的各种专题图, 也存在一定的属性与几何误差。

2 GIS 数据不确定性体系

根据国内外的研究现状, 邬伦等^[2]提出了 GIS 数据不确定性的框架体系(表 1)。该体系详细划分了 GIS 数据不确定性的各种来源和类型, 总结了各类 GIS 数据不确定性问题的主要研究内容及其在 GIS 当中的表现形式和所处的地位, 并对各种不确定性问题提出有针对性的研究方法^[3,4]。

3 GIS 数据不确定性理论与内容

3.1 位置不确定性建模理论

3.1.1 点状、线状地理目标位置不确定性建模

点状和线状地理目标的位置不确定性建模理论是矢量

GIS 位置不确定性的主要理论依据: 点元的位置不确定性理论是线状目标位置不确定性理论的基础; 而线状目标的位置不确定性理论又是矢量 GIS 中空间面与体的位置不确定性的理论基础。

空间实体位置不确定性基本模型包括点元、线元、曲线以及多边形(面元)模型。点元位置不确定性通常假定分布在以点元“真值”为中心的误差椭圆的邻域内。线元位置不确定性的建模则主要依据误差传播率、统计学和几何学等理论。在直线误差模型基础上又发展出曲线和多边形的误差模型。另外, 空间实体误差模型已发展成为评估更一般情形下实体的不确定性模型。

对于空间实体的位置误差分布早在多年前人们就已经做过研究。这种误差分布取决于数据获取方法等多种因素。最普通的数据获取方法是数字化。Bolsud 等^[5]、Jennifer 和 Walsby^[6]都发现数字化误差近似服从正态分布。随后 Meng 等^[7]和 Tong 等^[8]推断出数字化误差分别近似服从 NL 分布和最小 P 范分布。然而, 在处理诸如空间实体的自相关和互相关等问题时, 大多数的研究基于点元的位置不确定性服从正态分布的前提。

在点元的位置误差服从正态分布的假设前提下, 线元和多边形的位置误差分布主要采用模拟法和严密统计学理论两种方法来研究。Dutton^[9,10]采用模拟的方法描述了随机线元的误差分布, 认为构成线元端点的真实位置位于以端点点位均值为核心的误差区域内。如果点元仅有一个观测值, 那么此观测值即被认为是该点元的点位均值。基于点元正态分布假设的随机线元两端点模拟将获得一束线元族。这有助于探究出随机线元的误差分布特征。基于严密统计学^[11-14]的研究, 早期是在线元端点位置误差不相关和服从多维正态分布的假设前提下建立随机线元空间的位置不确定性模型。进一步地, 基于随机过程, Shi and Liu^[14]又提出了一种更一般性的用以研究线元位置误差分布的模型。该模型能够描述线元上点位相关时的各种误差分布情况。

误差带模型是包绕空间实体“真值”位置的一种带。误差带模型由 Pekal^[15]提出。Pekal 通过以沿线元的“真值”位置滚动的圆的集合来创建 ϵ -带。线元或多边形边界的位置不确定性可以用 ϵ -带的面积来描述。

Chrisman^[16]将 ϵ -带的理论模型进一步发展为实际应用模型。他将量测成果用于 ϵ -带, 建立美国地质调查局土地利用覆盖的 G IRAS 数字数据误差模型。在实用性方面, 改进的 ϵ -带宽建议由从数据获取第一步到最后一步累积起来的线元位置误差的统计函数来决定^[17]。误差带模型能够很容易地应用于空间操作中。然而, 由于假设线元的真实位置一定落入误差带之内, 因此, 误差带模型对线特征的形状很敏感。



作者简介: 邬伦(1964-), 男, 教授, 博士生导师, 主要从事 GIS 软件与技术、GIS 不确定性领域的研究工作。

E-mail: wulun@pku.edu.cn

收稿日期: 2006-07-24

表 1 GIS数据不确定性框架体系及主要研究内容

GIS数据的不确定性		主要研究内容		部分主要研究方法
GIS数据本身不确定性研究内容	位置数据的不确定性	数据源的不确定性	原始测量数据的不确定性 (GPS 全站仪、数字测图成果不确定性) 数字地图获取 (手扶跟踪数字化和扫描矢量化) 以及遥感数据的不确定性	1) 经典误差理论以及在此基础上的空间统计理论 2) 模糊集合理论 3) 灰色理论方法 4) 熵理论 5) 神经网络理论等方法 6) 模糊仿真方法
		GIS中基本实体不确定性	点实体的不确定性 线实体的不确定性 面实体的不确定性 体实体的不确定性	
		数据处理、空间分析过程不确定性	数据格式转换的不确定性 缓冲区分析不确定性 叠加分析不确定性 其他空间分析的不确定性	
	属性数据的不确定性	属性定义的不确定性	对现实世界定义不明确 模糊实体关系定义不确定	1) 目标模型和域模型 2) 概率论 3) 证据理论 4) 空间统计学 5) 模糊集合理论 6) 云理论 7) 粗集理论
		数据源的不确定性	实体边界界定不确定性 遥感数据边界提取不确定性	
		查询分析的不确定性	叠加分析的不确定性 查询过程中的不确定性 其他分析的不确定性	
	时域数据的不确定性	时间定义的不确定性	关于时间语义表达的模糊性带来的不确定性 界定事件开始与结束的不确定性	
	逻辑不确定性	属性一致性的不确定性 格式一致性的不确定性 拓扑关系一致性的不确定性		
数据的不完整性	几何数据的不完整性 属性数据的不完整性			
模型的不确定性	模型选择的不确定性 模型参数的不确定性 模型不确定性的传播与累加		1) 总体平均可靠性评价方法 2) 蒙特卡罗模拟	
不确定复合叠加	属性数据不确定性对位置数据不确定的影响 时间变更对属性数据不确定性的影响 时间变更对位置数据不确定性的影响 位置与属性数据不确定性叠加 时空不确定性叠加		缺陷率统计方法	
不确定性建模分析与表达	不确定性建模	数据不确定值的数据模型建立	科学计算可视化方法 GIS可视化方法	
	数据不确定性可视化表达	数据的不确定性的显示与表达 数据的不确定性可视化结果分析		
数据产品的不确定性评价	电子地图	电子地图不确定性评价	不确定性的叠加与传播	
	4D产品	DIG、DEM、DRG、DOQ产品的不确定性评价		
GIS工程不确定性评价与控制	GIS组件最优(不确定性最小)分割	GIS组件不确定性预测与评价 GIS组件间不确定性的叠加与传播 GIS组件最优分割模型	运筹学方法	
	GIS工程不确定性估计与控制	GIS工程不确定性预测与评价 GIS工程不确定性的控制方法与策略	建立预测模型软件工程方法、 项目管理方法	

Goodchild和Hunter^[18]提出的另一种确定带宽的方法,即通过计算落入 ϵ -带内的线元观测值比例来估计带宽,这种比例值应当不小于预设的容许度。从对线元形状敏感度方面讲,改进的 ϵ -带比以往的误差带模型更为适合。然而又可能存在这种错误情形,即一条线的两个不同的量测误差值对应着等宽的 ϵ 带。

除以上局限性外,还需要考虑到一条线元的两节点可能具有不同的分布。为了描述线元的位置不确定性,Shi^[11]提出了置信域误差模型。这里,线元端点的误差可能是具有不同参数的正态分布^[11-13]。这种置信域是一条围绕线元量测位置的带,线元的真实位置以不小于预设置信水平的概率被包含在该置信域内。

为了得到不随概率变化的误差带,范爱民和郭达志^[19]和李大军等^[20]利用信息论建立了一种称为误差熵的新型误差带-H带,即线元的真实位置位于该确定的区域内。根据范爱民和郭达志的误差带模型,H带带宽由熵系数决定,即由线元上点位误差熵与标准差的比率来决定或是由平均误差熵确定。熵系数仅当点元服从单峰正态分布时才能获得。然而,在点元服从坐标轴 x 向和 y 向具有不等标准差的二维正态分布情形下,熵系数无从确定。因此,李大军提出的H-带模型尚不用于处理点元误差分布不同时的一般性情形。

在基于线元端点误差无关的假设下,提出了两组模

型:置信域模型和线元误差分布模型。Shi^[11]提出二维随机线元置信域模型并进一步地发展为用于评估一维到 n -维任何空间线状实体的位置不确定性一般模型^[13]。这类模型的置信域是在基于点集中独立的位置误差的基础上导出的,其形状有别于 ϵ -模型-线元上任意点的位置不确定性都是不等的,这是对 ϵ -带的理论上的更进一步的发展,且较 ϵ -模型更趋于真实。另外,置信域模型使我们能够处理 n -维空间实体的位置不确定性这一更一般性的问题。

与以往误差带模型相比,G-带^[14]模型在建立线元位置不确定性模型方面提供了更为一般的解决方法。G-带模型可以用于建立两端点位置误差独立或者相关时的线目标位置不确定性模型。G-带模型考虑了线元是由线元上无数待定点集合而成的事实。当随机过程各向独立时,G-带模型退化为以往的误差带模型;当线目标各向同性且各向独立时,G-带模型退化为沿线元方向的平行带即 ϵ 带模型。因此,G-带模型提供了线元误差带建模的更一般性的方法。

GIS中曲线目标的表述方法比简单空间实体如点目标、直线目标、基于直线的多边形等复杂得多。Goodchild和Hunter^[18]提出了用于评估数字化线性目标位置不确定性的误差模型,并进一步应用于评估由多条折线逼近的海岸线的位置不确定性。Huang and Liu^[21]基于时间序列模式评估了等高线中的数字化误差,这类等高线均是以不规则曲线

模型来逼近的。史文中等^[22]提出了用于评估GIS中以规则曲线模型或不规则曲线模型逼近的曲线实体的位置不确定性的更一般方法。这种方法提出了如下几种不确定性描述指标:①以曲线法方向上均方根误差表示的 ϵ_s 模型;②以曲线上某点处最大误差表述的 ϵ_m 模型()。

汤仲安等^[23]定义了等概率密度误差模型的概念,提出了决定误差模型形状的因子与决定误差模型规模的尺度因子的概念,并给出了其相应的确定方法。汤仲安等提出了随机线“等概率密度误差模型的概念及其建模方法;史文中和汤仲安针对矢量GIS中不规则曲线的位置不确定性,研究了平面一般曲线的等概率密度误差模型。而平面规则曲线(如圆曲线、缓和曲线等)可以采用解析函数来表达,无需曲线函数方程的拟合过程,因而可以看做是平面不规则曲线位置不确定性误差模型研究的特例。

3.1.2 数字高程模型的不确定性

数字高程模型(DEM)是GIS赖以进行三维空间数据处理与地学分析的核心数据,其不确定性主要包括DEM数据的不确定性、DEM地形描述的不确定性和DEM数学分析的不确定性。

1) DEM数据的不确定性

DEM数据的不确定性主要表现在以下四个方面:①数据来源存在的不确定性:DEM是对自然真实地面的离散化采样过程,不确定性是采样过程中产生的系统错误的结果。从遥感相对的立体量测中获得的数据,即使经过校正,也可能存在随机和系统的残余误差。尤其是利用已知数字化地形图中高程信息产生高程点集,然后在已知点集上依据地形特征,构造内插函数,采用内插法获得所需间距的DEM数据也存在着不确定性特征;②数据结构引起的不确定性:为了反映高程数据的连续性特征,人们提出了许多数据结构,但是在地形平坦地区存在大量冗余数据,而且实际连续变化的地表面的离散化规则采样,必然引起地形描述的不确定,直接影响到实际应用效果。对于不规则三角网(TIN)数据结构,三角形的形状和大小,以及所表达的高程、坡度和坡向取决于局部地形表面的几何性质,这些三角形通常由一组Delaunay三角形构成,它们是从点的随机分布中产生的,所以具有不确定性特征;③内插法引起的不确定性:为了使DEM能表达连续曲面,只有高程值是不够的,还必须采用高程点间的内插法,因此,DEM的精度有限,存在不确定性;④DEM的重采样过程中引起的不确定性:即在保留了与原形地面较好逼近的情况下,将由原始资料派生的数据压缩或易于管理数据过程中进行了重采样。

2) DEM地形描述的不确定性

汤国安^[24]论证了DEM离散采样所造成的地形描述的不确定性,他在黄土高原地形进行了试验,证明了高程采样误差是影响DEM精度的重要因素,同时,即使DEM在所有高程采样点的误差均为零时,有限的DEM栅格采样点所构成的高程模型也只能是对真实地面的近似模拟,并对不同比例尺的DEM提取坡度等因子的不确定性进行了研究。

3) DEM的数学分析的不确定性

DEM分析模型的不确定性是DEM数据误差等多种因素耦合在一起综合影响的结果,其分析处理远比传统测量误差要复杂得多,其根源在于模型的影响。

3.2 属性不确定性

属性不确定性是指某一物体在GIS中被描述的属性与其真实的属性的差别,主要来源于空间现象的不稳定性、数据建模和分析过程中的误差。属性不确定性主要的研究成果有:基于高斯分布的分类数据不确定性自回归模型;基于知识推理的遥感数据解译不确定性模型;数据分类和边界划分不确定性模糊数学模型。通过与较高精度的数据集比较,用聚集的方法来控制遥感数据中的分类误差,从区域分类不确定性、边界定位误差和区域内部定量属性数据的抽样误差出发,综合进行属性数

据不确定性的度量和传播分析,以及基于缺陷率的统计模型用于评估GIS产品中属性不确定性。

对于离散属性数据插值模型精度问题的研究主要有:从遥感数据中提取DEM的不确定性;能够给出插值误差的Block-Kriging法;结合Monte Carlo法和Block-Kriging法研究栅格数据的误差评估和传播;权重矩阵和空间自相关统计的算法;用模糊数学方法研究边界分类和连续变量的不对称边界模型。

在GIS中,遥感数据是重要的数据源,除了DEM数据从遥感数据中获取外还有土地利用图的数据,包括林地耕地、城镇等数据也靠遥感数据解译而来而遥感数据是具有确定性和不确定性并存的,所以土地利用数据也可能存在着不确定性。

在GIS的属性数据除了从遥感数据中获得外,还来源于各种专题地图。不同的专题地图具有不同的属性可信度。大多数专题图数据,在不同程度上都有一定的属性误差,尤其对于大范围的专题图来说更是如此。不仅是根据地点少数实况监测站(如气象、水文、环境等)数据,按插值法绘制的各种专题图件,除了站所在位置的数据外,广大面上的数据多少存在不确定性,因为地面是平均的。对于运用模型,公式推导所得的数据,如果未经可信度验证的话,属性不确定性的可能性很大,即使经过了实况验证,也只能获可信度的大小,但也无法获误差的所在地。

目前人们正寻找更多的模型来研究属性不确定性问题,比如粗集理论、遗传BP算法、人工神经网络(ANN)、分形分析、层次递归算法、元胞自动机、云理论等等。

3.3 时域不确定性

时域不确定性是指在描述地理现象时,时间描述上的差错,数据的时间精度涉及数据的现势性,大多数采用元数据管理的方法。

如全国土地详查数据历时10年,开始得比较早的调查数据和开始得比较晚的调查数据,时间相差可达7~8年之久,而土地利用的状况,有些是固定10年,20年不变,但有一部分土地利用则每年都在变,因此,跨度达10年的土地利用数据,尤其是土地利用专题图数据,具有明显的不确定性。

还有遥感数据,对于每一幅遥感影像数据来说,无疑是同一瞬间的,但对于由多幅遥感影像拼接镶嵌起来的数据,则可能出现不同时间段的遥感数据,时间可以相差若干天、若干月和若干年。在同一数据集中,或同一幅专题图中,出现不同时段的数据就可能造成不确定性。

3.4 逻辑不一致性和数据不完整性

逻辑上的一致性指数结构内部的不一致性,尤其是指拓扑逻辑上的一致。在Euclidean空间,通常使用Egenhofer的相交模型结合距离函数来表达,数据的不完整性指对于给定的目标,GIS没有尽可能完全地表达该事物,这两个方面的问题仍待加强研究。

3.5 数据不确定性的传播

GIS中不确定性传播理论来源于误差分析理论、概率论、Taylor级数近似,许多自动误差传播系统都基于一定条件下的误差传播函数,如果叠置数据误差的方向一致,则以数据质量最差的层的质量指标来描述合并后的数据质量;如果不一致,则采用多重联合条件概率表示,这时参与误差传播计算的只是描述层的质量指标,有关各层内的误差变量并不参与计算。

GIS数据在输入、存贮、投影变换、格式变换、编辑、数据重组、数据处理、数据分析与分数、数据制图及输出过程中,都能产生误差或不确定性,这些不确定性还能随着数据流的过程传递,积累增强和放大,少数可能相互抵消,但误差的累计可能性大。

在数据格式转换过程中,误差是普遍存在的。从栅格数据到矢量数据的相互转换过程中,每一个模型计算所得的数据存在较大的差别,也难说谁对谁错。因此数据格式转存有明显的不确定性。另外,在多种数据叠加过程中,出现不符合的现象也是常有的,往往会出现新的“斑块”或双线条等问题,也无法采用那一条线合适,因此具有明

显的不确定性。

空间分析中的误差传播是近年来的一个研究热点,并取得了一定的进展。例如在空间叠加的误差传播建模、缓冲区分析、线状目标简化之误差传递。

3.6 空间数据不确定性管理

1998年美国地理信息和分析国家中心(NCGIA)首先提出了空间数据不确定性的管理问题。基于元数据的空间数据不确定性的管理方法受到了广泛的重视。运用元数据方法对空间数据的不确定性管理方法,已被采纳为国家标准。

关于空间数据不确定性管理是指在空间数据库中,确立定量和定性指标,对数据中所包含的不确定性状况进行区分、记录和管理。具体地说,就是区分空间数据不确定性的来源,根据其特点找出相应的不确定性指标,来对空间数据的不确定性进行分别表示和管理。其目的是避免空间数据不确定性的传播,从而确保数据的质量

Gan和 Shi提出了用元数据的方法管理空间数据质量,并且对数据质量的管理可以到特征级水平。空间数据不确定性管理的关键是如何确定空间数据不确定性的各类指标体系,以及在GIS数据库中的存贮和表达方式。现行的空间数据不确定性的描述,一般采用美国数字制图数据标准委员会的数据质量评定标准,包括空间数据的位置精度、属性精度、时态精度、逻辑一致性、完整性及数据志等。现行的指标体系适用于传统的矢量、栅格数据模型。

现行元数据方法中数据质量信息是面向数据集的,质量信息是文档性的说明,不能满足用户对数据不确定性进行分析的需要。

梅士员等^[25]提出基于特征的不确定性管理模型。特征是指对客观世界的高度抽象和全面表达。GIS不确定性数据的特征包含了属性、关系和功能等特性。不确定性将以特征属性的形式出现。对于特征空间数据的不确定性进行管理,可以分三个层次进行:①对不同特征类的不确定性进行分类管理;②对特征类实例的不确定性管理;③对基本几何点集所包含的不确定性的管理。三个层次对象中的不确定性指标具有不同的含义,包括定量和定性两类指标。特征GIS中空间数据不确定性的描述,必须考虑特征模型不同层次中的不确定性要素表达,以及不确定性在各层要素之间的传播关系。从技术实现的角度看,面向对象软件实现方法能满足基于特征的不确定性管理的需要,因此,该方法也能同时适应第四代GIS在理论和技术实现上的需要。

4 GIS数据不确定性处理的数学方法

就目前来说,GIS数据不确定性问题的研究理论和方法主要有模糊数学、数理统计、灰色系统理论、粗糙集、未确知数学、云理论、灵敏度分析等。

1) 模糊集

由于事物的复杂性,使其界限不分明,使其概念不能给出确定的描述,不能给出确定的评定标准,它向人们提供的宿信息称为模糊信息。现实世界中存在大量的模糊现象,如不同种类土壤类型的边界问题就是典型的模糊现象,再如森林和草原的过渡带中森林和草原的边界划分问题,污染水域与未污染水域的边界划分问题等都属于模糊性问题,需要借助模糊数学进行相应的研究。

模糊性是概念的外延不清晰,不能给出确定的评定标准。对某一事物,尽管我们不能说出它的明确程度,但可使用开区间 $(0, 1)$ 中的某个值表示事物的可信度。当越接近于1时,表明“同”的程度越来越大,当越来越靠近0时,表明“同”的程度越来越小,即“非”的程度越来越大。当等于1时,表明与人的认识无差异,即为“同”;当等于0时,表明与人的认识完全不同,即为“非”。在模糊集合中把 $[0, 1]$ 视为特殊信息归入其中,故模糊信息的值域为闭区间 $[0, 1]$,表示“同,异,非”三层含意。定义了模糊集合后,将其可以推广应用到经典数学的函数、矩阵、逻辑

群论、图论、测度、概率、拓扑等方面,从而得出了模糊函数、模糊矩阵、模糊逻辑、模糊群模图、模糊测度、模糊概率(模糊随机性)、模糊拓扑等分支,形成了模糊数学体系,可以解决许多模糊性问题。

2) 数理统计

现实世界客观存在的不确定性最为典型的为随机性,GIS是对现实世界的抽象和表达,因此GIS中的随机问题的研究也是比较重要的一部分。处理和解决随机性需要将概率论中的随机事件抽象为经典集合,根据集合的交、并、补、差等运算得出随机事件的积、和、对立、差等运算,从而求出各随机事件的概率并进而研究随机变量的分布、随机变量的数字特征以及大数定律和中心极限定理等重要问题。因此随机性还是以经典集合为基础的,可以用概率论进行分析和解决。

3) 粗糙集

粗糙集理论为处理具有模糊、不精确或不完全信息的分类问题提供了一种新的工具。其主要思想是在保持信息系统分类能力不变的前提下,通过知识约简,导出问题的决策或分类规则。粗糙集理论是一种处理模糊和不确定性问题的新型数学工具。粗糙集用上近似集和下近似集表示模糊边界线区域。由于上近似集和下近似集可以通过等价关系给出确定的数学公式进行描述,因此模糊区域中的含糊数目可以被计算出来,也即在真假二值之间的模糊程度可以计算。

4) 灰色系统理论

灰色概念是外延确定而内涵不确定的概念。灰性为不同于模糊性的另一类不确定性,它是部分已知部分未知的。由于事物的复杂性,信道上各种噪音的干扰以及接收系统能力的限制使得人类只能获得事物的部分信息或信息量的大致范围,而不能获得全部信息或确切信息。这种部分已知部分未知的宿信息称为灰色信息。具有灰色信息的不确定性为灰性,此类问题用经典集合无法解决和处理,但将经典集合推广为灰集合后就可以分析和处理上述问题。

灰色系统理论由我国学者邓聚龙在1982年所提出。灰色表示缺少、不完全、不确定等等现象。灰色系统是控制论观点和方法的延伸,它从系统的角度出发研究信息间的关系,即研究如何利用已知信息去揭示未知信息,也就是系统的“白化问题”。

在经典集合的特征函数表示法中,将特征函数拓宽为两个,分别记为 μ 称为上、下隶属函数。将值域 $\{0, 1\}$ 推广为 $[0, 1]$,则可得出灰集合定义。在灰集合的基础上,将其应用于经典数学中的函数、矩阵、群论、概率、拓扑等方面,则可以得出灰函数、灰矩阵、灰群、灰概率、灰拓扑等,形成灰色数学体系,可解决大量灰色不确定性问题。

5) 未确知数学

未确知性是由王光远教授提出的。未确知性是纯主观认识上的不确定性。它既无随机性又无模糊性,客观上是一种确定性事物,但决策者纯粹由于主观上对事物认识不清,该事物对决策者提供了一种不完整信息。我国学者吴和琴、刘开弟、王清印等在未确知性的研究方面做了大量工作,提出了基于未确知有理数的一套基本理论,初步建立了未确知数学的基本框架和体系。

6) 盲数

由于现实世界的复杂性,各种不确定性信息往往不是单独存在的,常常是多种不确定性的混合体,例如凡是有行为因素同时包含状态因素的任何体系中不确定性问题至少存在两种以上的不确定性,因为行为因素必然导致未确知性,而状态因素将导致随机性、模糊性、灰性或兼而有之。从信息的角度看,上述各种不确定性信息都是“单一”不确定性信息;而任意复杂的信息称为“信息混沌”,从信息混沌类中分离出一种最多同时具有随机信息、模糊信息、灰信息和未确知信息的较为复杂的信息称为盲信息,并用盲数进行表达和处理。

由盲数的定义,未确知有理数可以看做是盲数的特例。真盲数所包含的信息至少含有两种不确定性,因此可借助

盲数研究盲信息的数学表达和数学处理。对盲数的运算可采用类似于未确知有理数的运算方法。通过盲数的运算,可以得到盲数组合后的所有可能取值区间和取值可信度间的对应关系。用UM模型和BM模型可求算盲的可信度,供决策者从不同的角度进行分析。

7) 云理论

云理论是以研究定性定量间的不确定性转换为基础的,系统处理不确定性问题的一种新理论,包括云模型、虚云、云运算、云变换、不确定性推理等内容。其中云模型是定性定量间的不确定转换模型,它将概念的模糊性和随机性集成在一起。利用云模型可以有效地表示基于语言值的定性规则。云方法用期望值、熵和超熵这三个数字特征描述整个云团,实现定性和定量的转换,用于处理GIS中的模糊性和随机性为一体的属性不确定性问题。

8) 灵敏度分析

灵敏度分析用于定性或定量地评价模型参数误差对模型结果产生的影响,是模型参数化过程和模型校正过程中的有用工具。对于GIS中的模型不确定性的评价具有一定的指导意义。灵敏度分析也是模型参数校正过程中的一个非常有用的工具,其目的在于确定模型中哪些方面最容易在系统描述中引进不确定性。通过灵敏度分析可以确定模型各参数对输出结果影响的大小,在模型校正过程中重点考虑那些对输出结果影响大的参数,对于那些对模型结果几乎没有影响的参数可以不予考虑,这会在很大程度上减小模型校正的工作量。

9) 集对分析理论

集对是指具有一定联系的两个集合所组成的对子。集对分析认为,对这样两个集合的特性可以作同异反分析,并可定量地用联系度表达式来描述。

10) 集合预报法

集合预报是近年来国际上最新发展的新一代动力随机预报技术。集合预报的基本思想是,由于初始场误差、模式误差且大气具有高度非线性的混沌特性,使得数值预报结果具有不确定性,因而数值预报不应只是确定性的预报,应转变成具有动力学意义的概率预报。集合预报比单一预报效果好。

5 结束语

本文总结了当前GIS数据不确定性问题的研究进展,论述了GIS数据不确定性的框架体系,并详细讨论了GIS数据不确定性的核心理论和主要研究内容,最后对GIS数据不确定性的数学研究方法进行了归纳分析。除了数据不确定性外,其他GIS不确定性问题还需要进一步深入探讨,如数据处理、空间分析过程、数据产品、GIS工程中的不确定性问题等。

参考文献

- [1] 承继成, 郭华东, 史文中, 等. 遥感数据的不确定性问题 [M]. 科学出版社, 2004.
- [2] 邬伦, 于海龙, 高振纪, 承继成. GIS不确定性框架体系与数据不确定性研究方法 [J]. 地理学与国土研究, 2002, 18(4): 1-5.
- [3] WU Luan, YU Bo, ZHANG Yi. Modeling Fuzzy Spatial Objects Using Design Patterns [A]. In: Proceedings of 3th International Symposium on Spatial Data Quality [C], Vienna Austria, 2004.
- [4] GAO Zhen-ji, WU Lun, ZHOU Yue-yong, QIN Shi. Topological Relations between vague objects based on rough model [J]. Proceedings of 4th International Symposium on Spatial Data Quality [C], Beijing, China, 2005.
- [5] Bolstad P V, Gessler P, Lillesand T M. Positional error in manually digitized map data [J]. International Journal of Geographical Information Systems, 1990, (4).
- [6] Jennifer G. Walby. The Causes and Effects of Manual Digitizing on Error Creation in Data Input to GIS [A]. In: Em Innovations in GIS 2 [C], Edited by Peter Fisher Taylor and Francis, 1995, 113-122.

- [7] Meng X L, Shi W Z, Liu D J. Statistical tests of the distribution of errors in manually digitized cartographic lines [J]. Geographical Information Sciences, 1998, (4): 52-58.
- [8] Tong X H, Shi W Z, Liu D J. Error distribution error tests and processing of digitized data in GIS [A]. In: Accuracy 2000 Proceedings of the 4th International Symposium on Spatial Accuracy Assessment in Natural Resource & Environmental Sciences, Amsterdam, Cornet Books Inc, 2000.
- [9] Dutton G. A theory of cartographic error and its measurement in digital data base [A]. In: Proceedings of Auto-Carto 5 [D]. American Congress on Surveying and Mapping, Bethesda, 1992, 159-168.
- [10] Dutton G. Handling positional error in spatial databases [A]. In: Proceedings of the 5th International Symposium on Spatial Data Handling [C]. South Carolina, USA, 1992, 460-469.
- [11] Shi W Z. Modeling Positional and Thematic Error in Integration of GIS and Remote Sensing [Z]. ITC Publication, Enschede, 1994, 22-147.
- [12] Shi W Z. Statistical modeling uncertainties of three-dimensional GIS features [J]. Cartography and Geographical Information Systems, 1997, 24, 21-26.
- [13] Shi W Z. A generic statistical approach for modeling error of geometric features in GIS [J]. International Journal of Geographical Information Science, 1998, (12): 131-143.
- [14] Shi W Z, Liu W B. A stochastic process-based model for the positional error of line segments in GIS [J]. International Journal of Geographical Information Science, 2000, 14, 51-66.
- [15] Peckal J. On the Length of Empirical Curves. Ann Arbor in Michigan Inter-University Community of Mathematical Geographers (Discussion Paper 10) [Z], 1966.
- [16] Chrisman N R. A theory of cartographic error and its measurement in digital data base [A]. In: Proceedings of Auto-Carto 5 American Congress on Surveying and Mapping [C]. Bethesda MD, 1982, 159-168.
- [17] Chapman M A, Alsheikh A, Karim H A. Error modeling and management for data in GIS [A]. In: Proceedings of the CoastGIS97 [C]. Aberdeen, Scotland, 1997.
- [18] Goodchild M F, Hunter G J. A simple positional accuracy measure for linear features [J]. International Journal of Geographical Information Science, 1997, 11, 299-306.
- [19] 范爱民, 郭达志. 误差熵不确定带模型 [J]. 测绘学报, 2001, 30(1): 48-53.
- [20] 李大军, 龚健雅, 谢刚生, 杜道生. GIS中面元的误差熵模型 [J]. 测绘学报, 2003, 32, 31-35.
- [21] Huang Y G, Liu W B. Building the estimation model of digitizing error [J]. Photogrammetric Engineering and Remote Sensing, 1997, 63, 1203-1209.
- [22] 史文中, 童小华, 刘大杰. GIS中一般曲线的不确定性模型 [J]. 测绘学报, 2000, 29(1): 52-58.
- [23] 汤仲安, 王新洲, 纪现华. 矢量GIS平面随机线元误差模型建模机理 [J]. 武汉: 武汉大学学报, 2004, 29(11): 968-972.
- [24] 汤国安, 赵牡丹, 李天文, 等. DEM提取黄土高原地面坡度的不确定性 [J]. 地理学报, 2003, 58(6): 824-830.
- [25] 梅士员, 江南. 基于特征的空间数据不确定性管理模型 [J]. 遥感信息, 2004, 2: 42-45.

The spectrum characteristics of the radio frequency of Chinese first generation satellite navigation system — Beidou I

Abstract Beidou I is Chinese satellite navigation and positioning system of first generation. It covers territory and its surrounding area. The spectrum characteristics of the radio signal of "Beidou I" surveyed by foreign country are introduced in the paper. The construction and working frequency of "Beidou I" satellite navigation and position system as well as position precision of ground point to be located are discussed. At the same time it gives the discussion of the system's working program, the frequency spectrum of "Beidou" satellite as well as the feature map of "Beidou" satellite's signal.

Key words Beidou; navigation; spectrum

CHEN Jun-yong (State Bureau of Surveying & Mapping, Beijing 100830, China)

Data uncertainties in geographic information system

Abstract Based on an analysis of the status quo and accomplishment of research on GIS data uncertainty, the main frame and system of GIS data uncertainty is described primarily. Then the core theories and the main issues of GIS data uncertainty, especially the position uncertainty, the attribute uncertainty, the temporal uncertainty, the spread and management of uncertainty are discussed. At length, the mathematical methods of research on GIS data uncertainty are analyzed and concluded in detail.

Key words uncertainty; data uncertainty; GIS

WU Lun^①, CHENG Ji-cheng^①, SHI Wen-zhong^② (① Institute of RS and GIS, Peking University, Beijing 100871, China;

② Department of Land Surveying & Geo- Informatics, The Hong Kong Polytechnic University, Hong Kong)

Underwater positioning system based on DGPS

Abstract Underwater GPS positioning system includes four parts which are GPS differential reference station, GPS buoy subsystem, underwater pinger, ship-based data processing and control center. GPS differential station provides real-time differential correction data, and GPS buoy receives differential correction data to calculate real-time dynamic position and output real-time buoy's space coordinate. At the same time GPS buoy also gets the hydro- phonic signal to measure the distance to the underwater targets. The ship-based data processing and control center is the most important part of the system. All the data are transmitted to this center and the underwater targets' space coordinates are resolved in this center. In this paper, the system structure and each part's function are introduced firstly. Then, the time measuring and positioning algorithm are developed. Some experiments and numerical results are showed in the end.

Key words underwater positioning system; DGPS; GPS buoy; hyperbola

WANG Quan^①, CHENG Peng-fei^①, ZHANG Chuan-yin^①, WANG Ze-min^②, CAI Yan-hui^①, DING Jian^① (① Chinese Academy of Surveying and Mapping, 100039; ② No. 715 Research Institute of Shipbuilding Industry Corp. of China, 310012)

Regularization error estimation of spherical harmonic coefficients from SGG observation

Abstract Gravity gradient is the second order derivative of gravitational potential, which can be observed by satellite gradiometer. The spherical harmonic coefficients of the gravity field can be obtained from gravity gradients through regularization scheme. Based on the analysis of the regularization scheme, a method to estimate the error of regularization solution is presented in this paper. The results of this paper will provide technical materials for Chinese future satellite gravity mission.

Key words earth gravity field; satellite gravity gradiometer; GOCE; regularization; error estimation

CHANG Xiao-tao^{①②}, ZHANG Chuan-yin^②, DING Jian^{①②} (① Laboratory of Dynamical Geodesy, Chinese Academy of Sciences, Wuhan

China 430077; ② Chinese academy of surveying and mapping, Beijing 100039, China)

Study on the anti- solution of latitude function

Abstract The elements of earth-ellipsoid such as radius M of curvature in meridian are the function of latitude. According to the definition of function and anti-function, the authors demonstrate the existence of anti-solution and deduce the anti-solution formula of elementary function. They give the anti-solution computing program of transcendental equation for CASIO fx-4800p calculator. The correctness of all anti-solution program have been proofed by example computation.

Key words function of latitude; elementary function; transcendental equation; anti-solution program

ZHONG Ye-xun^{①②}, HUANG Jun-hua^{①③}, LI Zhan-yuan^{①②} (① Guangxi Regional Bureau of Surveying and Mapping, Nanning, 530003, China; ② School of Resource & Environment, Guangxi Teachers' College, Nanning, 530001, China; ③ Department of Civil Engineering, Guilin University of Technology, Guilin, 541004, China)

Automated cycle-slip detection and correction for GPS triple-frequency undifferenced observables

Abstract In the process of GPS triple frequency observables, the rms of the integer ambiguity isn't fit for detecting and correcting cycle slips for the pseudo-range noise. The noise will be reduced with a simple running average filter or a process of combination for the raw observables. It is recommended that selecting observables after a running average filter or combination as a detection observable can detect and correct cycle slips for single frequency. Four principles are proposed for constructing detection observables and three geometry-free linear combinations are chosen according to these principles. Cycle slips can be detected and corrected for triple-frequency undifferenced observables with these selected detection observables.

Key words cycle slip; ambiguity resolution; detection observables; geometry-free phase combination; MW combination.

FAN Jian-jun, WANG Fei-xue, GUO Gui-rong (Satellite Navigation R&D Center, School of Electronic Science and Engineering, National University of Defence Technology, Changsha, Hunan, China, 410073)

Robust Kalman filtering for the marine sounding data

Abstract After analyzing the pattern of sounding, the models based on the algorithms of Kalman filtering were presented in this paper to filter the survey-boat positions and depths along a sounding line. However, the existing methods could not remove the alias soundings in fact. So, a filtering method with a truncated weight function is designed to remove outliers, such as gross errors, false echo. The method has been tested using observed data. The results show that the robust Kalman filtering with the truncated function is really robust in resisting the outliers in soundings.

Key words marine sounding; a sounding line; robust Kalman filtering

LIMing-san^{①②③}, LIU Yan-chun^{①②}, BAO Jing-yang^{①②}, LV Zhi-ping^③, XIAO Fu-min^{①②} (① Department of Hydrography & Cartography, Dalian Naval Academy, Dalian, 116018, China; ② the Geomatics and Applications Laboratory, Liaoning Technical University, Fuxin, 123000, China; ③ Institute of Surveying and Mapping, Information Engineering University, Zhengzhou, 450052, China)

The precision analysis of satellite stereophotogrammetry using physical method to determine exterior orientation elements

Abstract This paper analyses the precision of satellite stereophotogrammetry in the case of using physical method to determine exterior orientation elements. It infers theoretical precision calculating formula of satellite stereophotogrammetry and calculates the theoretical precision of SPOT5 HRS. Meanwhile it makes a test with few GCPs. The experiment